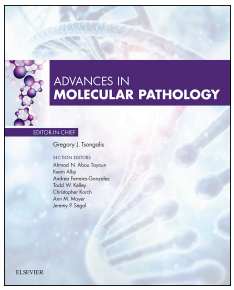
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal Pre-proof



The Genomic Landscape of SARS-CoV-2: Surveillance of *Variants of Concern*

M. Shaheen S. Malick, Helen Fernandes

PII: S2589-4080(21)00006-5

DOI: https://doi.org/10.1016/j.yamp.2021.06.006

Reference: YAMP 71

To appear in: *Advances in Molecular Pathology*

Please cite this article as: Malick MSS, Fernandes H, The Genomic Landscape of SARS-CoV-2: Surveillance of *Variants of Concern*, *Advances in Molecular Pathology* (2021), doi: https://doi.org/10.1016/j.yamp.2021.06.006.

The Genomic Landscape of SARS-CoV-2: Surveillance of *Variants of Concern*

M. Shaheen S. Malick and Helen Fernandes

Department of Pathology & Cell Biology,

Columbia University Irving Medical Center,

630 West 168th Street, New York, NY, USA

Corresponding Author: Helen Fernandes, hf2340@cumc.columbia.edu

Key words:

Genomic Surveillance, SARS-CoV-2 mutations/variants, Variants of Concern, COVID-19 Pandemic

Key points:

- SARS CO-V2 variants/strains may have varying degrees of transmissibility, virulence, pathogenicity and response to vaccines.
- Monitoring and identifying variants is crucial to control the current pandemic.
- Vaccines developed for the wildtype virus may have reduced efficacy against some variants of concern.
- Global surveillance is important for public health initiatives.

Abstract:

Novel mutations that drive the evolution of SARS-CoV-2 variants, are constantly emerging. Given the acquired mutations of some of these variants compared to wild-type virus, identifying and monitoring variants through sequence-based surveillance is essential to control their spread. Prompt stratification of new variants into level of concern based on the mutational signatures, is important for understanding their response to current therapeutics and vaccines. Variants of concern have increased levels of infectivity, pathogenicity and transmissibility that enable the spread of the virus globally.

The Genomic Landscape of SARS-CoV-2: Surveillance of *Variants of Concern*

Epidemiological investigations are extremely helpful for understanding the impact of these variants on the COVID-19 pandemic. Moreover, in light of recent surges in India and Brazil, by the Delta and Gamma variants respectively despite high levels of seroprevalance , current mathematical models for herd immunity may need to be re-evaluated.

As we continue to sequence SARS-CoV-2 genomes, novel mutations that drive viral evolution are constantly emerging. Such mutations generate a rich diversity of viral strains and lineages with varying degrees of transmissibility, virulence and pathogenicity. As these variants evolve, it is essential to monitor the efficacy of current diagnostics, therapeutics, and vaccines in efforts to control their spread.

Mutations are random events that naturally occur during replication of the error-prone viral genome. The host-pathogen interaction drives the selective pressure by which variants capable of escaping aspects of the natural and, perhaps, induced immunity subsequently start to dominate the circulating pool of infecting strains.[1] Rapid characterization and active monitoring of emerging variants to understand their potential impact on SARS-CoV-2 countermeasures is critical to the control of the pandemic. The Center for Disease Control and Prevention (cdc.gov) has classified emerging variants into 3 categories based on the degree of impact of COVID in the US - Variants of Interest; Variants of Concern; Variants of High Consequence. (See https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html for further details). Variants of concern have been in part responsible for the waves of infection that we see globally (Figure 1).

There are several key reasons why genomic surveillance has now become crucial in the global efforts to manage the COVID-19 pandemic. Since late fall of 2020, four new SARS CoV-2

variants of concern emerged; the United Kingdom (UK) variant, known as 501Y.V1, or B.1.1.7; the South African variant known as 501Y.V2 or B.1.351; the 501Y.V3 or P.1 variant from Brazil and most recently the B.1.617.2 variant in India. These variants share some common mutations but each has their own unique genomic landscape. On May 31st 2021, the WHO (World Health organization) elected to simplify the names of theses VOC with Greek alphabets. The four variants of concern are now Alpha, Beta, Gamma and Delta respectively (https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/).

Some variants appear to have enhanced transmissibility or infectivity than the wild-type virus and other circulating strains. These acquired characteristics of the SARS CoV2 virus have been responsible for the continued surge of cases and the increased fatality in different parts of the globe (Figure 2). Additionally, early clinical trials have demonstrated reduced vaccine efficacy against novel strains compared to original strain, most notable being the 501Y.V2 (B1.351) variant in South Africa. As several nations enter into a second wave of cases, reports of reinfection with the new circulating variants arise regularly, generating concern on the status of acquired immunity from prior exposure.

Variants of concern in resource limited settings:

India and Brazil represent two important locations to examine given the recent surge in cases primarily due to novel variants of concern. The B.1.617 variant first detected in India in October 2020 is characterized by L452R, D614G and P681R substitutions within the spike protein domain. Three sub-lineages have been characterized, B.1.617.2 being the predominant strain currently devastating India during its second wave. Based on the observed growth in sequenced

cases, the B.1.617.2 is more transmissible than the B.1.1.7 variants. [2 3]. An additional cause for concern as India battles this variant is despite high levels of seroprevalence, indicating prior infection, in regions such as Delhi (56%), Mumbai (up to 75%) and Hyderabad (54%) the number of cases in these cities is once again surging.[4] This pattern is not at all unique to India.

In Manaus, Brazil, a variant with an E484K substitution was associated with a recent surge in cases despite a reported seroprevalence of 76% as of October 2020. [5] An important caveat to seroprevalence studies is that detectable humoral immunity does not capture the entire scope of immunity, including T cell reactivity. However, the estimated infection rate in Manaus is above the theoretical herd immunity threshold of 67% using a case reproduction number (R_0) of 3. [6] The abrupt surge in hospitalizations that hit Manaus in January 2021, followed a period during which physical distancing requirements were eased (see ref. 5). This surge that surprisingly included SARS-CoV-2 reinfection of individuals in addition to new individuals, was attributed to a novel strain containing an E484K substitution.[7] Reinfection in the city of Manaus has been associated with the P.1 lineage characterized by ten unique spike protein mutations, including E484K and N501K. [8] In one Brazilian case of re-infection 9 months after initial infection, the re-infecting virus carried all lineage-defining mutations of P.1 but 11 additional amino acid substitutions in the S-protein relative to the primary infecting virus.

Interestingly, the unique molecular signatures that characterize the novel strains appear to be responsible for the lack of protection in previously infected individuals. Variants with E484Q or E484K substitutions have been found to have reduced neutralization by protective antibodies, or are, in other words, able to escape optimal protection by humoral immunity. This has been demonstrated effectively by an international research collaboration led by the Africa Health

Research Institute (Durban, South Africa). This team employed *in vitro* live-viral neutralization assays, with plasma from individuals infected during the country's first and second waves. They clearly demonstrated that convalescent plasma from individuals infected with wild-type virus during the first wave, was unable to neutralize the 501Y.V2 (B1.1.351) virus that was responsible for roughly 97% of cases during the second wave in South Africa.[9]

Seroprevalence and efficacy of vaccines:

Returning to the idea of herd immunity, the resurgence in Brazil and India despite a seroprevalance of >70% causes us to reconsider our current mathematical models for herd immunity. This model is represented by the equation $R = (1 - p_c)(1 - p_I)R_0$, where R is the effective reproduction number, p_c, is the relative reduction in transmission due to non-pharmaceutical measures, p_I is the proportion of immune individuals, and R_0 is the reproduction number in the absence of controls in a fully susceptible population. In this model, herd immunity is achieved when $R < 1$, hopefully indicating an end to future large outbreaks. If this model holds true then the only way to explain the surges in Brazil and India is if current seroprevalance studies no longer accurately represent the proportion of immune individuals, or p_I, since seropositivity of an older strain would not represent immunity to a novel variant of concern. The other potential explanation would be if the R_0 of these variants were greater than that of the prior circulating strains. Potentially both of these may be true. Whatever the explanation, given the dynamic nature of the virus, genomic surveillance is key to understanding the variant landscape of the infecting virus.

Notably, we have not yet raised the question of whether current SARS-CoV-2 vaccines are effective against novel strains, or whether they affect p_c in the way they have increased p_c for

early strains. Preliminary results from clinical trials suggest vaccines developed for wildtype virus, may have reduced efficacy against variants of concern. So far, in South Africa, at least two clinical trials during the second wave reported a reduction in vaccine efficacy against the 501Y.V2 strain. The NVX-CoV2373 subunit vaccine (Novavax) showed a decrease in efficacy from 89.3% to 49.4%.[10] Importantly, this trial reported no differences in infection frequency between SARS-CoV-2-seropositive and SARS-CoV-2-seronegative participants in the placebo arm suggesting that infection with variants other than 501Y.V2 does not protect against re-infection with 501Y.V2. Rollout of the AstraZeneca ChAdOx1 AZD1222 chimpanzee adenovirus-vectored vaccine is currently paused in South Africa after displaying only 10% efficacy against the 501Y.V2 variant, compared with an efficacy of 75% against earlier variants in South Africa.[11]

Role of genomic surveillance for monitoring a pandemic:

Representative, quality, timely and continuous genetic surveillance of SARS-CoV-2 is critical to the COVID-19 outbreak response. In addition to acquiring information on the infectivity and virulence of the virus, quick identification genomic alterations helps to elucidate the molecular mechanisms that allow variants of concern to evade vaccine-induced immunity and/or targeted therapy. Sequencing can also potentially alert us to variants that may eventually render current diagnostic tests ineffective.

Currently there is a lack of capacity for genomic surveillance globally. Rapid progress is occurring in scaling up genomic surveillance. However, the level of surveillance has not reached the necessary breadth to track and manage the pandemic. Large numbers of samples were sequenced in the UK, the USA, Australia, and Iceland and data were shared promptly. However,

as of May 22, 2021, the U.S. has shared 11.7 sequences per 1000 cases and currently ranks 25th in the number of genomic sequences shared to the online genome repository GISAID compared to the United Kingdom with 85.1 sequences shared per 1000 cases and 16 days to deposition.[12] Since November 2020, the U.S. CDC has accepted SARS-CoV-2 samples from state health departments and other public health agencies for sequencing and further necessary characterization in a program called the National SARS-CoV-2 Strain Surveillance System (NS3) (see ref. 1). While this program continues to grow there is additional work to be done, especially as the U.S. has the highest number of reported cases. The lack of capacity for testing alone, in the U.S. does not explain the deficiency in sequencing. There is an enormous virus-sequencing capacity, but funding, coordination and systemic problems in sharing samples and data are hurdles that need to be addressed. A study published on medRxiv preprint server, has predicted that sequencing at least 5% of all cases is necessary to detect emerging variants. As of May 22, the last 7-day average reported cases was 27,789, meaning that at 750 samples per week we are below the recommended 5%. Further strengthening of the sequencing capacity at a global level would help in the fight against not only the current pandemic but also future outbreaks of viral diseases. Low- and middle-income countries some of which are even now struggling to find adequate numbers of vaccines, have an even greater deficiency in genomic surveillance. [13]

The public-health mantra of the pandemic since the outset has been 'test, trace, and isolate.' As the phylogenetic tree of SARS-CoV-2 blossoms it is now abundantly clear that genomic surveillance must buttress our testing. Otherwise, we may be unsuccessful in tackling this to the best of our capacity. Perhaps the greatest aspect of human creativity is the ability to do highly impactful work with limited resources, making the best of what we have available. With

diagnostic stewardship, it is possible to scale up genomic surveillance worldwide, though we will need to be strategic about how the scarce genomic sequencing resources are utilized and shared globally. In areas where this infrastructure is limited, genomic surveillance still can and must be done, though it will need to be much more targeted. For example, in resource limited settings genomic surveillance may be limited to only patients who present with reinfections. This can then be broadened to entire geographical areas that develop surges in infection despite having reached herd immunity, or regions with unexpectedly high rates of transmission. It is also important to correlate variant and lineage details with epidemiological data to inform public health decisions. We have entered a new era in the COVID-19 pandemic. However, it may also be a promising turning point in pandemic preparedness.

References:

[1]Genomic Surveillance for SARS-CoV-2 Variants – CDC. May 17, 2021. Available at: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance.html. Accessed May 21, 2021.

[2] SAGE 89 minutes: Coronavirus (COVID-19) response, 13 May 2021. Eighty-ninth SAGE meeting on COVID-19, 13 May 2021. Available at: https://www.gov.uk/government/publications/sage-89-minutes-coronavirus-covid-19-response-13-may-2021/sage-89-minutes-coronavirus-covid-19-response-13-may-2021. Accessed May 26, 2021.

[3] Narang, D. COVID-19: The Second Wave May Not Be the Last – but Which One Will Be? Science, The Wire. April 22, 2021. Available at: https://science.thewire.in/health/covid-19-the-second-wave-may-not-be-the-last-but-which-one-will-be/ . Accessed May 24, 2021.

[4] Hyderabad SARS-CoV-2 Serosurvey – Ministry of Science and Technology, Govt. of India, Council of Scientific and Industrial Research, Centre for Cellular and Molecular Biology. Available at: https://www.ccmb.res.in/presscovrg/Hyderabad_SARS_CoV2_serosuvey.pdf. Accessed May 26, 2021.

[5] Sabino EC, Buss LF, Carvalho MPS, Prete CA Jr, Crispim MAE, Fraiji NA, Pereira RHM, Parag KV, da Silva Peixoto P, Kraemer MUG, Oikawa MK, Salomon T, Cucunuba ZM, Castro MC, de Souza Santos AA, Nascimento VH, Pereira HS, Ferguson NM, Pybus OG, Kucharski A, Busch MP, Dye C, Faria NR. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. Lancet. 2021 Feb 6;397(10273):452-455. doi: 10.1016/S0140-6736(21)00183-5.

[6] Fontanet A, Cauchemez S. COVID-19 herd immunity: where are we? Nat Rev Immunol. 2020 Oct;20(10):583-584. doi: 10.1038/s41577-020-00451-5.

[7] Nonaka CKV, Franco MM, Gräf T, de Lorenzo Barcia CA, de Ávila Mendonça RN, de Sousa KAF, Neiva LMC, Fosenca V, Mendes AVA, de Aguiar RS, Giovanetti M, de Freitas Souza BS. Genomic Evidence of SARS-CoV-2 Reinfection Involving E484K Spike Mutation, Brazil. Emerg Infect Dis. 2021 May;27(5):1522-1524. doi: 10.3201/eid2705.210191.

[8] Faria NR, Claro IM, Candido D, Franco LAM, Andrade PS, Thais M, Silva CAM, Sales FC, Erika R, Aguiar RS, et al. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. VirologicalOrg. 2021;

[9] Cele S, Gazy I, Jackson L, Hwa SH, Tegally H, Lustig G, Giandhari J, Pillay S, Wilkinson E, Naidoo Y, Karim F, Ganga Y, Khan K, Bernstein M, Balazs AB, Gosnell BI, Hanekom W, Moosa MS; Network for Genomic Surveillance in South Africa; COMMIT-KZN Team, Lessells RJ, de Oliveira T, Sigal A. Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. Nature. 2021 May;593(7857):142-146. doi: 10.1038/s41586-021-03471-w.

[10] Shinde V, Bhikha S, Hoosain Z, Archary M, Bhorat Q, Fairlie L, Lalloo U, Masilela MSL, Moodley D, Hanley S, Fouche L, Louw C, Tameris M, Singh N, Goga A, Dheda K, Grobbelaar C, Kruger G, Carrim-Ganey N, Baillie V, de Oliveira T, Lombard Koen A, Lombaard JJ, Mngqibisa R, Bhorat AE, Benadé G, Lalloo N, Pitsi A, Vollgraaff PL, Luabeya A, Esmail A, Petrick FG, Oommen-Jose A, Foulkes S, Ahmed K, Thombrayil A, Fries L, Cloney-Clark S, Zhu M, Bennett C, Albert G, Faust E, Plested JS, Robertson A, Neal S, Cho I, Glenn GM, Dubovsky F, Madhi SA; 2019nCoV-501 Study Group. Efficacy of NVX-CoV2373 Covid-19 Vaccine against the B.1.351 Variant. N Engl J Med. 2021 May 20;384(20):1899-1909. doi: 10.1056/NEJMoa2103055

[11] Madhi SA, Baillie V, Cutland CL, Voysey M, Koen AL, Fairlie L, Padayachee SD, Dheda K, Barnabas SL, Bhorat QE, Briner C, Kwatra G, Ahmed K, Aley P, Bhikha S, Bhiman JN, Bhorat AE, du Plessis J, Esmail A, Groenewald M, Horne E, Hwa SH, Jose A, Lambe T, Laubscher M, Malahleha M, Masenya M, Masilela M, McKenzie S, Molapo K, Moultrie A, Oelofse S, Patel F, Pillay S, Rhead S, Rodel H, Rossouw L, Taoushanis C, Tegally H, Thombrayil A, van Eck S, Wibmer CK, Durham NM, Kelly EJ, Villafana TL, Gilbert S, Pollard AJ, de Oliveira T, Moore PL, Sigal A, Izu A; NGS-SA Group; Wits-VIDA COVID Group. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. N Engl J Med. 2021 May 20;384(20):1885-1898. doi: 10.1056/NEJMoa2102214.

[12] Global sequencing coverage, COVID-19 CoV Genetics, Enabled by data from GISAID. Available at: https://covidcg.org/?tab=global_sequencing. Accessed May 23, 2021.

[13] Cyranoski D. Alarming COVID variants show vital role of genomic surveillance. Nature. 2021 Jan;589(7842):337-338. doi: 10.1038/d41586-021-00065-4. PMID: 33452508.

Figure legends

Figure 1: Depicts the global pattern of COVID-19 infections and fatalities from April 2020 to April 2021. The waves represent variants of concern that emerge in different countries. *Source: World Health Organization*

Figure 2: Shows the global landscape of SARS CoV-2 infections over time. The patterns of emergence, representing increased infectivity is seen. *Source: World Health Organization*